

# 《逸周書》舊版造字轉碼說明

中研院資訊所文獻處理實驗室  
中央研究院語言所文獻語料小組  
2010/07/06 丁玟伶

轉碼主要工作是把檔案中的舊版造字轉換成 Windows XP 能支援的 Unicode 字形，Unicode 目前共收錄漢字 70194 個字，而 XP 只能支援 20902 個字（詳如表一），不支援之字將以構字式表達。例：造字編號 996 的「衡」字，Unicode 編碼是 27602，由於 XP 並不支援，仍需使用構字式「行△童」。

表一、Unicode 的字數及編碼區段

Unicode	新增字數	新增編碼區段	總字數	WinXP
1.1 版	20902	4E00-9FFF	20902	支援
3.0 版	6582	3400-4DFF	27484	不支援
3.1 版	42710	20000-2A6D6	70194	不支援

## 一、舊版造字轉碼分析：

《逸周書》使用舊版造字 46 個，字頻 176 次，這 46 個造字中，39 個可轉成 Windows XP 能顯示的字，字頻 159 次；另外 7 個字必須轉成構字式，字頻 17 次。

轉碼完成製作轉碼分析表，請參考附件一《逸周書》轉碼分析表，欄位說明如下：

- (一) 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
- (二) 造字：舊版造字。
- (三) 頻次：舊版造字在文件的出現次數。
- (四) Big5：造字的 Big5 碼。
- (五) Unicode：造字所對應的 Unicode 碼。

- (六) WinXP：造字在 Windows XP 的對應字形。
- (七) 構字式：Windows XP 無法對應字形改用構字式。
- (八) 備註凡例：備註欄中記錄轉碼後字形及修改原因，凡例如下：
1. 異體字問題：為了使用者查詢和使用的方便，在處理異體字時最主要的方式是以標準字取代，除非是專有名詞或特殊情形，如：人名、地名等。例：造字編號 4134 的「衛」字，是「衛」的異體字，以標準字「衛」取代。
  2. 錯字，以程式全部取代：與原書字形不符，並查詢教育部異體字字典確認非異體字後，皆歸類為錯字，以程式取代為正確字形。例：檔案中使用造字編號 5505 的「耆」字，而原書使用之字為「耆」，查詢教育部異體字字典確認「耆」並非「耆」之異體字，因此歸為錯字，以程式取代為正確字形「耆」。
  3. Unicode 字型呈現差異：Unicode 字型與舊版造字有些微差異，但只是字體風格差異，實際上仍為同一個字，因此仍取 Unicode 字型。如編號 3985 的「璫」字，Unicode 字型呈現為「璫」，實際上仍為同一字。
  4. 待造字：Unicode 及漢字構形資料庫皆未收錄的舊漢籍造字，正在等待補造字中，所以「造字」欄空白無法看到字形。如編號 3065 的「𠂇𠂈」。

附件一、《逸周書》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
828	𠂇	1	8E6A	8DD4	𠂇		
899	𠂈	1	8ED3	99DE	𠂈		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
979	蕙	1	8F64	56CF	蕙		
995	獾	1	8F74	734F	獾		
996	衝	1	8F75	27602		行△童	
998	潰	1	8F77	3D52		汜△費	
1383	糠	2	91E0	7A45	糠		
1530	紉	1	92D6	7CFA	紉		
1565	緜	2	92F9	7DDC	緜		
1893	袞	1	9548	886E	袞		
1981	讐	5	95C2	8B90	讐		
2816	備	43	9AF4	50C3	備		
3064	鯛	1	9CB2	9C02	鯛		
3065		8	9CB3			一△彡	待造字。
3787	冢	5	8152	51A2	冢		
3855	孽	2	81B8	5B7C	孽		
3883	徧	2	81D4	5FA7	徧		
3929	黎	1	8243	68C3	黎		
3965	裁	1	8267	70D6	裁		
3976	犁	1	8272	7282	犁		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3985	璫	1	827B	7447	璫		Unicode 字型 呈現差異。
4014	蓋	5	82BA	8462	蓋		
4038	竈	2	82D2	7AC8	竈		
4068	綫	1	82F0	7DAB	綫		
4080	羣	27	82FC	7FA3	羣		
4105	菓	1	8356	83D3	菓		
4134	衛	11	8373	885E	衛		衛，異體字問 題。
4156	諂	1	83AB	8B1F	諂		
4184	迹	3	83C7	8FF9	迹		
4190	邳	3	83CD	90C4	邳		
4242	鬪	6	8442	9B2D	鬪		
4267	鼈	9	845B	9F08	鼈		
4733		4	8756			形丘-4 □	
4756	呂	1	876D	382F	呂		
4793	祿	3	87B4	7958	祿		Unicode 字型 呈現差異。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4902	兗	1	8862	5156	兗		
4904	尅	6	8864	5C05	尅		
4955	龜	1	88B9			龜	
5034	弃	1	8949	5F03	弃		Unicode 字型 呈現差異。
5126	暮	1	89C7	8B29	暮		
5189	宀	1	8A47	5B82	宀		
5289	慤	1	8ACD	6164	慤		
5302	昏	2	8ADA	662C	昏		
5461	𠂔	1	8BDC	6B6B	𠂔		
5505	𠂔	1	8C49	264BE		老会占	耆，錯字，以 程式全部取 代。
5687	𠂔	1	8D62	4AC9		𠂔	

## 二、手動取代表說明：

《逸周書》檔案中出現「●」字形，共計21字，比對原書後以Unicode字形或構字式手動取代，並製作「●」字手動取代表，請參考附件二。欄位說明如下：

1. 頁碼：「●」字形所在位置之原書頁碼。

2. 檔案原文摘錄：摘錄《逸周書》檔案有「●」字形的文句。
3. 原書字形：「●」在原書中之字形，以此字取代檔案原文的「●」。
4. 備註：記錄修改相同詞句的次數或其他事項。

附件二、《逸周書》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 1927	治之初●	麻△虫	
p. 1940	不卯不●	足△巽	
p. 1953	遂●集之于	走△秦	
p. 1955	則●●乃不和	扌△舞 扌△舞	
p. 1958	以●伯父	佺	
p. 1964	不●曰戴	寔△心	
p. 1965	●、過也	寔△心	
p. 1972	青馬黑●	𠂔△駢	
p. 1972	東越海●	合△虫	
p. 1973	會稽以●	單△黽	
p. 1973	卜盧以●牛	纟△九	
p. 1973	●牛者	纟△九	
p. 1973	夷用●木	門△焦	
p. 1975	●鯛之醬	魚△吳	
p. 1975	禁豐一●	角△豪	
p. 1975	侯。樂，鈇●	王△業	

頁碼	檔案原文摘錄	原書字形	備註
p. 1990	始●竈	𠂔𠂔曾	待造字
p. 1991	歲將●終	儂	
p. 1991	●雨	雨𠂔眾	待造字
p. 1993	味辛而不●	膠	